

*STIMULUS DISPARITY AND PUNISHER CONTROL OF HUMAN  
SIGNAL-DETECTION PERFORMANCE*

CELIA LIE AND BRENT ALSOP

UNIVERSITY OF OTAGO, NEW ZEALAND

The present experiment examined the effects of varying stimulus disparity and relative punisher frequencies on signal detection by humans. Participants were placed into one of two groups. Group 3 participants were presented with 1:3 and 3:1 punisher frequency ratios, while Group 11 participants were presented with 1:11 and 11:1 punisher frequency ratios. For both groups, stimulus disparity was varied across three levels (low, medium, high) for each punisher ratio. In all conditions, correct responses were intermittently reinforced (1:1 reinforcer frequency ratio). Participants were mostly biased away from the more punished alternative, with more extreme response biases found for Group 11 participants compared to Group 3. For both groups, estimates of discriminability increased systematically across the three disparity levels and were unaffected by the punisher ratios. Likewise, estimates of response bias and sensitivity to the punisher ratios were unaffected by changes in discriminability, supporting the assumption of parameter invariance in the Davison and Tustin (1978) model of signal detection. Overall, the present experiment found no relation between stimulus control and punisher control, and provided further evidence for similar but opposite effects of punishers to reinforcers in signal-detection procedures.

*Key words:* signal detection, punishment, stimulus disparity, response bias, point-loss, mouse-click, human

The study of detection focuses on decision-making in situations of uncertainty. In everyday life, many situations require organisms to make choices involving the detection or identification of stimuli. For example, an animal must decide whether a plant is toxic or safe to eat, a pedestrian must decide whether it is safe to cross the road, or a quality control officer must decide whether a product meets certain production standards. In these situations, negative outcomes arising from incorrect choices and positive outcomes arising from correct choices both play an integral part in decision-making.

The signal-detection task is a discrete-trial procedure, where, on each trial, the subject is presented with one of two discriminative stimuli ( $S_1$  or  $S_2$ ). These can vary on some physical (e.g., intensity) or temporal (e.g., stimulus presentation duration) dimension, or can be the presence or absence of a stimulus (e.g., a sound against background

noise). The subject then makes one of two available responses ( $B_1$  or  $B_2$ ) to identify which stimulus was presented. With the combination of two stimulus types and two response alternatives, four response outcomes are possible: responding  $B_1$  following  $S_1$  ( $B_{11}$ ), responding  $B_2$  following  $S_1$  ( $B_{12}$ ), responding  $B_1$  following  $S_2$  ( $B_{21}$ ), and responding  $B_2$  following  $S_2$  ( $B_{22}$ ). Correct responses ( $B_{11}$  and  $B_{22}$ ) can be reinforced, for example, with money (Johnstone & Alsop, 2000), food (McCarthy & Davison, 1979), or brain stimulation (Terman, 1970). Errors ( $B_{12}$  and  $B_{21}$ ) usually have no consequence; however, they can also be punished (e.g., time-out; Hume & Irwin, 1974).

Traditionally, models of signal detection have investigated performance as a function of two independent parameters: discriminability and response bias. Discriminability measures how well the subject can tell the two stimuli ( $S_1$  and  $S_2$ ) apart. Response bias measures the subject's tendency to make more of one response over another response, irrespective of which stimulus ( $S_1$  or  $S_2$ ) was presented. Response bias is often manipulated by varying the relative frequency or magnitude of reinforcement for  $B_{11}$  and  $B_{22}$  responses. Models of signal detection aim to provide measures of discriminability which are independent of changes in response bias, and

---

This research was conducted as part of the first author's doctoral thesis, supported by a University of Otago Postgraduate Scholarship. Portions of these data were presented at the 30<sup>th</sup> Annual Conference of the Society for the Quantitative Analyses of Behavior in San Diego, California.

Address correspondence to Celia Lie, Department of Psychology, P.O. Box 56, Dunedin, New Zealand (e-mail: celia@psy.otago.ac.nz).

doi: 10.1901/jeab.2010.93-185

measures of response bias which are unaffected by changes in discriminability (see Macmillan & Creelman, 2005).

Davison and Tustin (1978) proposed a model of signal-detection performance based on the generalized matching law (Baum, 1974). The generalized matching law can be written (in logarithmic form) as:

$$\log\left(\frac{B_1}{B_2}\right) = a \log\left(\frac{R_1}{R_2}\right) + \log c \quad (1)$$

where  $B_1$  and  $B_2$  denote the number of responses made on Alternatives 1 and 2 respectively, and  $R_1$  and  $R_2$  denote the number of reinforcers obtained by making  $B_1$  and  $B_2$  responses. The parameter  $c$  (or  $\log c$ ) is a measure of inherent bias independent of changes in the reinforcer ratio (e.g., side or color preference). The parameter  $a$  (termed sensitivity) measures the effect that variations in the reinforcer distribution ( $R_1/R_2$ ) have on the distribution of the subject's response distribution ( $B_1/B_2$ ).

Davison and Tustin (1978) stated that when two stimuli ( $S_1$  and  $S_2$ ) are indistinguishable in a signal-detection procedure, behavior follows the generalized matching law (Equation 1). When the stimuli are more distinguishable, the subject makes more correct ( $B_{11}$  and  $B_{22}$ ) responses and Davison and Tustin formulated two separate equations to describe this relation. In the case where reinforcers ( $R_{11}$  and  $R_{22}$ ) are intermittently obtained for correct responses, choice in detection tasks is described on  $S_1$  trials by:

$$\log\left(\frac{B_{11}}{B_{12}}\right) = a \log\left(\frac{R_{11}}{R_{22}}\right) + \log c + \log d \quad (2)$$

and on  $S_2$  trials by:

$$\log\left(\frac{B_{21}}{B_{22}}\right) = a \log\left(\frac{R_{11}}{R_{22}}\right) + \log c - \log d. \quad (3)$$

All notations are as above, and the parameter  $\log d$  measures stimulus discriminability. When  $\log d = 0$ ,  $S_1$  and  $S_2$  are indiscriminable and Equations 2 and 3 reduce to the generalized matching law. As  $S_1$  and  $S_2$  become more disparate, the subject makes more correct responses ( $B_{11}$  and  $B_{22}$ ); thus,  $\log d$  is additive in Equation 2 and subtractive in Equation 3.

Algebraic subtraction of Equations 2 and 3 provides a bias-free measure of discriminability:

$$\log d = 0.5 \log\left(\frac{B_{11}B_{22}}{B_{12}B_{21}}\right) \quad (4)$$

where all notations are as above. Equation 4 predicts that discriminability should be independent of the reinforcer ratio ( $R_{11}:R_{22}$ ). However, Johnstone and Alsop (1999) re-analyzed a number of past detection studies and found greater estimates of  $\log d$  for unequal reinforcer ratios than equal reinforcer ratios.

Algebraic addition of Equations 2 and 3 provides a discriminability-free measure of response bias. This bias is attributable to the effects of both the distribution of reinforcers and any inherent bias ( $\log c$ ):

$$\begin{aligned} \log b &= 0.5 \log\left(\frac{B_{11}B_{21}}{B_{12}B_{22}}\right) \\ &= a \log\left(\frac{R_{11}}{R_{22}}\right) + \log c \end{aligned} \quad (5)$$

where all notations are as above. Equation 5 states that response bias,  $\log b$ , should follow the generalized matching law and this expression includes no term for discriminability. Thus, it predicts that the effects of the reinforcer ratio (and the sensitivity of the subject's behavior to changes to that ratio,  $a$ ) should be independent of changes in discriminability ( $\log d$ ; McCarthy & Davison, 1980). This prediction, however, has received mixed support with some studies finding no systematic relation and others finding an inverse relation between sensitivity and discriminability. As an example, Figure 1 plots the relation between estimates of sensitivity and discriminability for three studies that varied both the reinforcer frequency ratio and the disparity between  $S_1$  and  $S_2$  (Alsop & Davison, 1991; Alsop & Porritt, 2006; McCarthy & Davison, 1984). These studies employed controlled reinforcer procedures where the obtained reinforcer ratios were constrained to match the arranged reinforcer ratios (Stubbs & Pliskoff, 1969). Because subjects made too few errors ( $B_{12}$  and  $B_{21}$  responses) at the highest discriminability levels to calculate accurate estimates of  $a$  and  $\log d$ , only

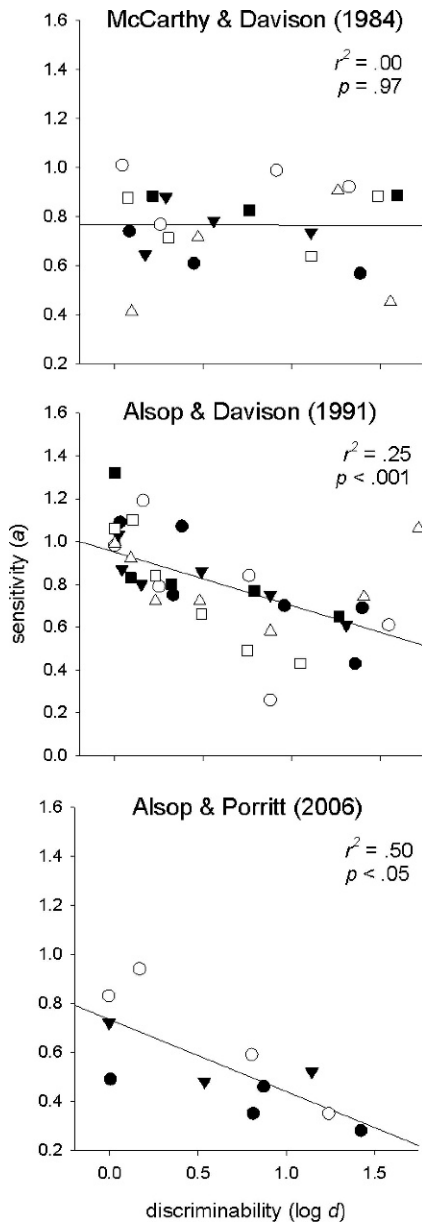


Fig. 1. Sensitivity of response bias to changes in reinforcer ratio (a) plotted against estimates of discriminability (log d) for individual subjects from three previous studies: McCarthy and Davison (1984—top), Alsop and Davison (1991—middle), and Alsop and Porritt (2006—bottom). See text for details.

conditions with estimates of log d below 1.75 (i.e., around 98% correct assuming no inherent bias) were plotted.

Figure 1 (top) shows the data from McCarthy and Davison (1984), which arranged four

stimulus discriminability levels, and for each level, arranged three different reinforcer frequency ratios (4:1, 1:1, and 1:4). For example, in the 4:1 condition, pigeons obtained reinforcement for correct B<sub>11</sub> responses around four times more frequently than for correct B<sub>22</sub> responses. McCarthy and Davison found no systematic relation between sensitivity to the reinforcer ratios and discriminability. Figure 1 (middle) plots the data from Alsop and Davison (1991), who also arranged a number of different stimulus discriminability levels with three reinforcer frequency ratios (9:1, 1:1, and 1:9). Unlike McCarthy and Davison however, they found an inverse relation between sensitivity and discriminability—as their pigeons' estimates of discriminability increased, estimates of sensitivity to the reinforcer ratios systematically decreased. Finally, Figure 1 (bottom) plots the data from Alsop and Porritt (2006), where instead of varying the reinforcer frequency ratios, they varied the reinforcer magnitude ratios (3:1, 1:1, and 1:3). For example, in the 3:1 condition, pigeons received 6-s access to food for reinforced B<sub>11</sub> responses and only 2-s access to food for reinforced B<sub>22</sub> responses. Similar to Alsop and Davison's finding with reinforcer frequency, Alsop and Porritt found an inverse relation between sensitivity to reinforcer magnitude and discriminability.

The mixed findings regarding the relation between sensitivity to reinforcement and discriminability are not limited to the above studies, but these provide an illustration of the conflicting outcomes. Of particular relevance to the present study, Johnstone and Alsop (2000) conducted the only study investigating the effects of stimulus disparity and reinforcer frequency with human participants. In their study, one group of participants received a 4:1 controlled reinforcer ratio across conditions while another group received a 1:4 controlled reinforcer ratio. Participants in both groups completed four conditions where the disparity between S<sub>1</sub> and S<sub>2</sub> stimuli was varied. Figure 2 plots the results from their study. As expected, when stimulus disparity increased, discriminability (log d) also increased (Figure 2, top). Furthermore, estimates of response bias (log b—Figure 2, bottom) were negative for the 1:4 group (black bars) and positive for the 4:1 group (grey bars), demonstrating reinforcer control (i.e.,

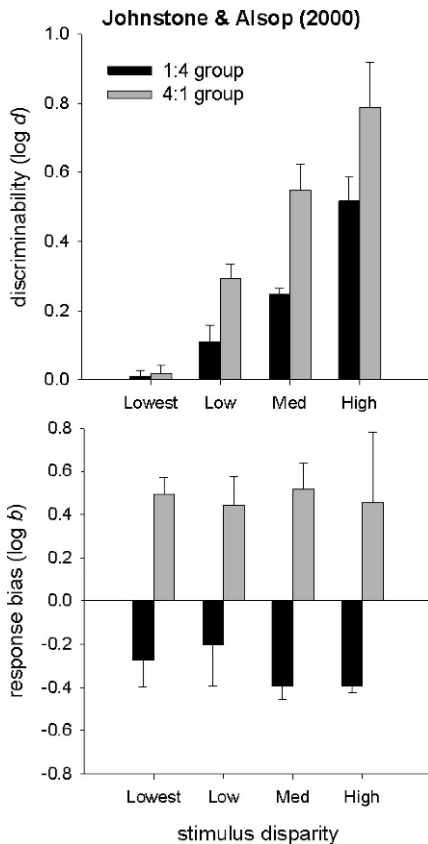


Fig. 2. Mean discriminability ( $\log d$ —top) and response bias ( $\log b$ —bottom) for each stimulus disparity level calculated across participants in the 1:4 reinforcer ratio group (black bars) and 4:1 reinforcer ratio group (grey bars) in Johnstone and Alsop's (2000) study.

participants were biased towards responding on the alternative associated with the greater frequency of reinforcement). However, there were no systematic effects of stimulus disparity on response bias for either group. These data are consistent with the notion of an independence between reinforcer control ( $\log b$  or  $a$ ) and stimulus control ( $\log d$ ).

The studies described above varied the relative frequencies or magnitudes of reinforcers for the two types of correct responses. In fact, most detection research (using both nonhuman and human animals) has focused on the effects of positive outcomes for correct responses. In contrast, little attention has been given to the effects of negative outcomes for errors. This lack of research on aversive consequences is of concern because many real-life detection scenarios involve both pos-

itive outcomes (e.g., crossing the road safely) and negative outcomes (e.g., getting hit by a car). Recently, Lie and Alsop (2009) found that punishers had parallel but opposite effects to reinforcers using a signal-detection task with human participants. In one of their experiments, participants occasionally received points for correct responses and lost points for errors. The punisher frequency ratio was varied across four conditions (5:1, 2:1, 1:2, and 1:5) and Lie and Alsop found that participants were systematically biased away from the response alternative associated with the higher rate of punishment.

The present experiment used the same detection task as Lie and Alsop (2009) to investigate the relation between relative punisher frequency and stimulus disparity. There were two groups of participants. For each group, the relative frequency of punishment ( $P_{21}:P_{12}$ ) was varied across two levels (1:3 and 3:1, Group 3; 1:11 and 11:1, Group 11) and across three levels of stimulus disparity. These results were compared to Johnstone and Alsop's (2000) study which found no relation between the effects of stimulus disparity and the relative frequency of reinforcers.

## METHOD

### Participants

Twenty-four undergraduate students from the University of Otago were recruited from the psychology participant pool. Group 3 consisted of 12 females aged between 18 and 21 years ( $M = 19.5$  years). Group 11 consisted of 11 females and 1 male aged between 20 and 24 yr ( $M = 21.0$  yr).

### Apparatus

The experiment was conducted in a small room (2.3 m  $\times$  3.0 m), where the window blinds were closed to minimize visual distractions. The participant sat facing a PC computer, with his or her head approximately 0.5 m from a 43-cm (17-inch) color monitor. The computer presented the instructions, ran the signal-detection program, and recorded the participant's responses. The program was written in VisualBasic 6.0. Stimuli were 12  $\times$  12 arrays presented in the centre of the white screen, with each array position occupied by either a blue or red alien cartoon character ("greeblie") measuring approximately 8 mm

wide and 9 mm high (see Figure 3, Lie & Alsop, 2009). Stimuli with more blue objects than red objects were classed as “more blue” ( $S_1$ ) and stimuli with more red objects than blue objects were classed as “more red” ( $S_2$ ). The difference between the number of blue and red objects in the array varied across three disparity levels. At the easiest level (high stimulus disparity), there were 77 of one color and 67 of the other color. At the medium level (medium disparity), there were 75 of one color and 69 of the other color. At the hardest level (low disparity), there were 73 of one color and 71 of the other color. For a particular disparity level, the program randomly determined the arrangement of blue and red greeblies within the stimulus array on each trial. Participants responded on a two-key response panel (with telegraph Morse keys) connected to the computer’s USB port via a Lab Jack<sup>TM</sup> interface device, with the left key labeled as the response for “more blue” ( $B_1$ ) and the right key labeled as the response for “more red” ( $B_2$ ).

#### Procedure

The first 12 participants were assigned to Group 3 and the next 12 participants were assigned to Group 11. Participants in each Group completed six experimental sessions (one condition per session) no less than 24 hr apart and no more than one week apart. Before the start of the first session, participants read an information sheet and signed an informed consent form. Each session consisted of a set of instructions presented on the computer screen followed by the experimental trials. The following instructions were presented:

Screen 1: *This is a simple computer game. You will see some patterns of blue greeblies and red greeblies. You must decide if there are more blue ones or red ones, and then press the blue button or red button. Here is an example of a pattern:*

Screen 2: *If there are more blue greeblies, press the blue button.* An example array showing more blue greeblies was presented.

Screen 3: *If there are more red greeblies, press the red button.* An example array showing more red greeblies was presented.

Screen 4: *Try to be as accurate as possible. On most trials you will receive no feedback. You could be correct or incorrect. Sometimes, when you are correct, you will gain one point. Sometimes, when you are incorrect you will lose*

*one point. When you get to 50 points, the session will end and you can go.*

Screen 5: *As you go, a red bar (like that on the right) will show you how close you are to finishing the experiment. When the red bar gets to the top, you can go! Any questions? If not, you are ready to start the session.* A vertical bar was presented on the right side of the screen.

If the participant had no questions, the experimenter started the trials. Each trial began with a 1-s presentation of a small animated picture of a juggler in the middle of the screen (warning stimulus). A stimulus array containing either more blue or red greeblies then appeared and remained on the screen until the participant responded on the response panel, or for a maximum of 3 s. The stimulus presentation probability was .5 for all conditions; that is, participants were equally likely to be presented with  $S_1$  (more blue greeblies) or  $S_2$  (more red greeblies). If the participant had not responded after the 3-s stimulus presentation, the screen remained blank until the participant pressed a response key.

Following each response, there were three possible consequences. First, if neither a reinforcer nor a punisher was scheduled, the screen went blank for 1 s; that is, no feedback was given. This was followed by a 1-s intertrial interval (ITI) with a blank screen, and the next trial began.

Second, if the participant made a correct response (i.e.,  $B_{11}$  or  $B_{22}$ ) and a reinforcer was scheduled for that response, the center of the screen displayed the statement: “Correct! You are one point closer to finishing the session.” The start of this presentation was accompanied by a 1-s “ta da!” sound. If the participant made a correct “more blue” response (i.e., the participant responded on the left key to a “more blue” stimulus— $B_{11}$ ), a picture of a blue tick (✓) appeared in the lower left corner of the screen. Likewise, if it was a correct “more red” response (the participant responded on the right key to a “more red” stimulus— $B_{22}$ ), a picture of a red tick appeared in the lower right corner of the screen. The vertical thermometer bar also went up one “space” (out of 50 spaces). The statement, colored tick, and thermometer bar remained on the screen for 3 s. This was followed by a 1-s ITI and the next trial began.

Finally, if the participant made an incorrect response and a punisher was scheduled for



that response, then the center of the screen displayed the statement: “*Incorrect! You are one point further from finishing the session!*” The start of this presentation was accompanied by a 1-s “argh!” sound. If the participant made an incorrect “more blue” response (i.e., the participant responded on the left key to a “more red” stimulus— $B_{21}$ ), a picture of a blue “x” appeared in the lower left corner of the screen. If the participant made an incorrect “more red” response ( $B_{12}$ ), a picture of a red “x” appeared in the lower right corner of the screen. The vertical thermometer bar went down one space. The statement, colored “x”, and thermometer bar remained on the screen for 3 s. This was followed by a 1-s ITI and the next trial began.

The session ended when the participant reached a net total of 50 points. If the participant had not reached 50 points within

50 min, the session also ended. At the end of the last trial, the screen displayed the statement: “*Congratulations. You have reached the end of the session*” for 4 s and the program ended.

The six experimental conditions varied stimulus disparity and the distribution of punishers for errors. Stimulus disparity was varied across three levels, high, medium, and low, as described above. Incorrect responses were intermittently punished and the distribution of punishers ( $P_{21}$ : $P_{12}$ ) varied across two levels at each level of stimulus disparity. For Group 3 participants, the punisher distributions were 3:1 and 1:3; for example, if the punisher distribution was 3:1, participants were three times more likely to receive  $P_{21}$  punishers than  $P_{12}$  punishers. For Group 11 participants, the punisher distributions were 11:1 and 1:11. Correct responses were intermittently reinforced and the distribution of

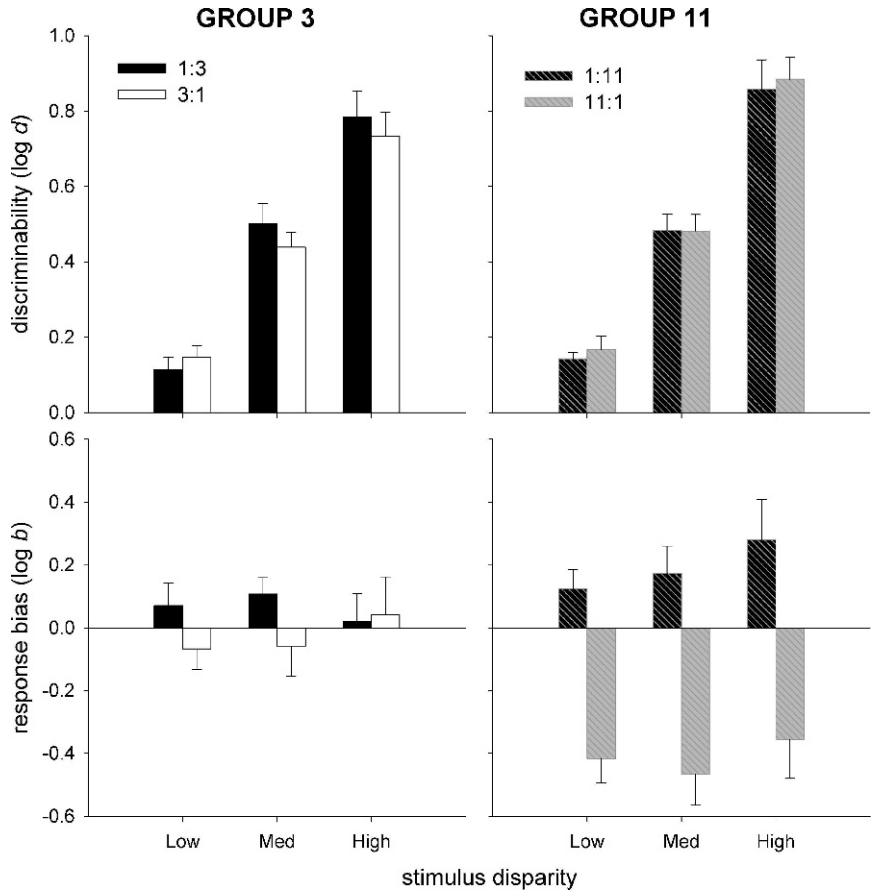


Fig. 3. Mean discriminability ( $\log d$ —top) and response bias ( $\log b$ —bottom) for each stimulus disparity level calculated across all participants in Group 3 (left) and Group 11 (right).

reinforcers (point gains) was held constant and equal (1:1) across all conditions; that is, participants were equally likely to receive  $R_{11}$  and  $R_{22}$  reinforcers for correct  $B_{11}$  and  $B_{22}$  responses, respectively. The presentation order of the conditions was partially counterbalanced across all the participants for each group, with the constraint that conditions with the same punisher ratio were run consecutively (see Appendix for condition order).

The distributions of reinforcers and punishers were arranged using interdependent scheduling (e.g., McCarthy & Davison, 1984; Stubbs & Pliskoff, 1969) where the computer program randomly selected the next correct response ( $B_{11}$  or  $B_{22}$ ) to be reinforced, or the next incorrect response ( $B_{21}$  or  $B_{12}$ ) to be punished, in accordance with the arranged reinforcer and punisher ratios. The reinforcer (or punisher) then had to be received before the program selected the next response to be reinforced or punished. The rates of reinforcement and punishment differed depending on disparity level to try to control for a potential confound where the overall rates of reinforcement and punishment could vary as a function of stimulus disparity. For example, at the highest disparity level, where participants are correct more often (and thus incorrect less often), participants could obtain reinforcers at a higher overall rate, and punishers at a lower overall rate, compared to lower disparity levels (where participants make more errors). For the high disparity conditions, the overall rate of reinforcement was based on a VI 15-s schedule and the overall rate of punishment was based on a VI 10-s schedule. For the medium disparity conditions, reinforcement was based on a VI 10-s schedule and punishment was based on a VI 15-s schedule. For the low disparity conditions, reinforcement was based on a VI 10-s schedule and punishment was based on a VI 40-s schedule.

The numbers of obtained reinforcers ( $R_{11}$  and  $R_{22}$ ) and punishers ( $P_{21}$  and  $P_{12}$ ) were calculated across all trials for each session. The numbers of left key ("more blue") responses following  $S_1$  ( $B_{11}$ ) and  $S_2$  ( $B_{21}$ ), and right key ("more red") responses following  $S_1$  ( $B_{12}$ ) and  $S_2$  ( $B_{22}$ ) were calculated across the last 120 trials from each condition for each participant. Trials before the last 120 trials were discarded to allow participants adequate contact with the reinforcers and punishers.

## RESULTS

The Appendix shows the results for individual participants in each group. These data show that the manipulations were successful at keeping the reinforcer ratios constant and equal across conditions. Furthermore, the obtained punisher ratios approximated the arranged punisher ratios for each condition.

Individual estimates of discriminability ( $\log d$ ) and response bias ( $\log b$ ) were calculated for each participant in each condition using Equations 4 and 5 (Davison & Tustin, 1978), respectively. Because there were a few instances where participants made zero responses in the last 120 trials for a particular response type ( $B_{11}$ ,  $B_{12}$ ,  $B_{21}$ , or  $B_{22}$ ), a correction of 0.25 was applied to all response counts for  $\log b$  and  $\log d$  calculations (Brown & White, 2005). These estimates were then averaged across all participants in each group for each condition. Figure 3 presents these mean estimates of discriminability ( $\log d$ —top) and response bias ( $\log b$ —bottom) for Group 3 (left) and Group 11 (right) for the three stimulus disparity levels and the different punisher ratios (3:1 and 1:3 for Group 3; 11:1 and 1:11 for Group 11). Figure 3 (top) shows that, as expected, estimates of discriminability significantly increased as stimulus disparity increased for both Group 3,  $F(2,22) = 94.45$ ,  $p < .001$ , and Group 11,  $F(2,22) = 141.7$ ,  $p < .001$ . Furthermore, there were no significant differences in discriminability between the two punisher ratios for Group 3,  $F(1,11) = .791$ ,  $p = .39$ , or Group 11,  $F(1,11) = .229$ ,  $p = .64$ .

Figure 3 (bottom) shows that mean response biases were more extreme for Group 11 (right) compared to Group 3 (left). The effect of punisher ratio on response bias was significant for both Group 3,  $F(1,11) = 4.895$ ,  $p < .05$ , and Group 11,  $F(1,11) = 21.77$ ,  $p < .01$ . Furthermore, there were no significant differences in response bias across the three stimulus disparity levels for Group 3,  $F(2,22) = .168$ ,  $p = .85$ , or Group 11,  $F(2,22) = .748$ ,  $p = .49$ . Despite not finding a significant effect, however, there appeared to be an increasing trend for the 1:11 punisher ratio across the three disparity levels for Group 11 (Figure 3, bottom right—black striped bars). To investigate this further, Figure 4 presents the individual response bias data for Group 11 participants. Figure 4 (left) shows that the

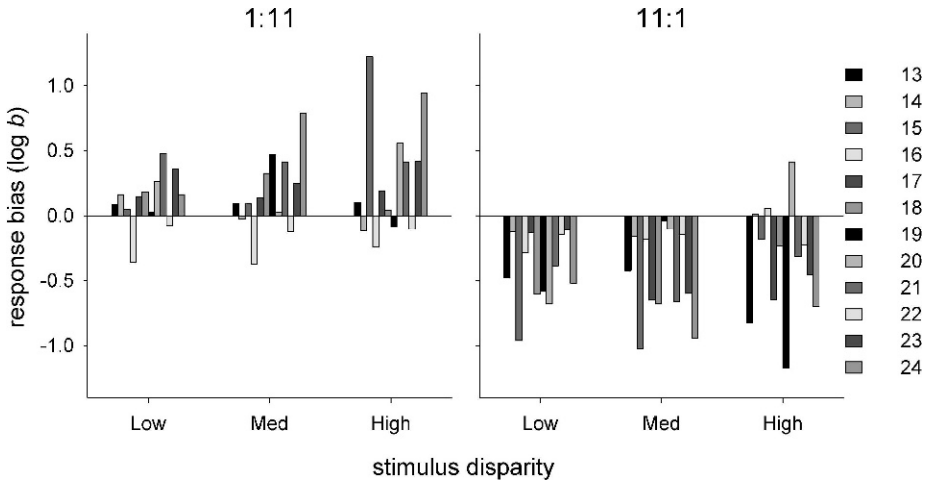


Fig. 4. Individual estimates of response bias ( $\log b$ ) for the 1:11 (left) and 11:1 (right) arranged punisher ratios at each stimulus disparity level for participants in Group 11.

apparent increase in mean response bias estimates across the stimulus disparity levels for the 1:11 condition was predominantly due to extreme data points from 2 participants (Participants 15 and 24). Figure 4 (right) confirms the pattern seen in Figure 3 (bottom right—grey striped bars), with no effect of disparity level on response bias estimates in the 11:1 condition.

Another way to assess the effects of punisher ratios at each level of stimulus disparity is to compare the change in bias estimates from the 3:1 (Group 3) and 11:1 (Group 11) conditions with their reversals (i.e., 1:3 and 1:11 respectively). This was done using Equation 6 for each participant at each disparity level:

$$\begin{aligned} \log b &= 0.5 \log \left( \frac{B_{11}B_{21}}{B_{12}B_{22}} \right) \\ &= a_p \log \left( \frac{P_{12}}{P_{21}} \right) + \log c \end{aligned} \quad (6)$$

where all notations are as above. Equation 6 is analogous to Equation 5 in that it measures the sensitivity of behavior to the changes in the punisher ratios ( $a_p$ ). There are, of course, only two conditions for each fit of Equation 6, therefore the estimates of  $a_p$  will not be very precise. However, it does allow for a rough comparison between the present experiment and previous studies which examined changes in sensitivity of behavior to changes in reinfor-

er ratios ( $a$ , Equation 5) across different levels of stimulus disparity. In the present experiment, positive  $a_p$  estimates indicate a systematic bias away from the more frequently punished alternative (i.e., towards the less frequently punished alternative). Because there were 8 sessions (out of 144) where some participants received no punishers for one response alternative, a correction of 0.25 was applied to all punisher counts for  $\log (P_{12}/P_{21})$  calculations.

Table 1 displays individuals' estimates of sensitivity to changes in the punisher ratio ( $a_p$ ) for the three different levels of disparity in Group 3 and Group 11. For Group 3, positive  $a_p$  values (indicating a bias away from the more punished alternative) were found for 10 out of the 12 participants at the lowest level of disparity ( $M = 0.14$ ). At the medium disparity level, positive  $a_p$  values were found for 8 out of the 12 participants ( $M = 0.18$ ). However, at the highest disparity level, a less consistent pattern was found, with 6 participants obtaining negative  $a_p$  values (indicating a bias towards the more punished alternative), and the remaining 6 participants with positive  $a_p$  values ( $M = -0.04$ ). One-sample  $t$ -tests conducted on the  $a_p$  values for each disparity level found that estimates of sensitivity were significantly greater than zero for the lowest disparity level only; Low:  $t(11) = 2.223$ ,  $p < .05$ ; Medium:  $t(11) = 2.008$ ,  $p = .07$ ; High:  $t(11) = -0.298$ ,  $p = .77$ . A repeated-measures ANOVA found no significant differences in



Table 1

Sensitivity to the punisher ratio ( $a_p$ ) for each participant in Group 3 and Group 11 at each stimulus disparity level.

Part.	Disparity level		
	Low	Med	High
<b>GROUP 3</b>			
1	0.36	0.39	-0.22
2	0.30	-0.22	0.18
3	0.06	0.05	0.23
4	-0.34	0.63	0.33
5	0.20	-0.04	0.49
6	0.40	0.63	-0.77
7	0.04	-0.09	-0.46
8	0.03	-0.20	0.35
9	0.38	0.19	-0.16
10	0.14	0.61	0.44
11	-0.08	0.01	-0.66
12	0.17	0.25	-0.20
<b>GROUP 11</b>			
13	0.27	0.26	0.39
14	0.13	0.06	-0.06
15	0.49	0.48	0.79
16	-0.04	-0.09	-0.15
17	0.13	0.38	0.41
18	0.36	0.49	0.14
19	0.30	0.25	0.70
20	0.46	0.06	0.07
21	0.42	0.55	0.40
22	0.04	0.01	0.07
23	0.23	0.47	0.42
24	0.34	0.91	0.82

sensitivity estimates between the three disparity levels,  $F(2,22) = 1.306$ ,  $p = .29$ .

For Group 11, almost all participants obtained positive  $a_p$  values at all three disparity levels (Low:  $M = 0.26$ ; Medium:  $M = 0.32$ ;

High:  $M = 0.33$ ), therefore showing some sensitivity to change in the punisher ratio. Only 2 participants (Participant 16 at all three levels, and Participant 14 at the highest disparity level) had small negative  $a_p$  values. One-sample  $t$ -tests found that  $a_p$  values were significantly greater than zero for all three disparity levels, Low:  $t(11) = 5.385$ ,  $p < .001$ , Medium:  $t(11) = 3.888$ ,  $p < .01$ , High:  $t(11) = 3.548$ ,  $p < .01$ . Consistent with the findings from Group 3, no significant differences in sensitivity estimates were found across the three disparity levels for Group 11 participants,  $F(2,22) = .617$ ,  $p = .55$ .

Figure 5 plots the estimates of sensitivity to changes in the punisher ratio ( $a_p$ ) against estimates of discriminability ( $\log d$ ). For Group 3 (left), estimates of sensitivity appeared lower at higher discriminability values compared to lower discriminability values, but this correlation was not significant ( $r = -.23$ ,  $n = 36$ ,  $p = .18$ , two-tailed). For Group 11 (right), sensitivities appeared greater at higher discriminability levels than lower discriminability levels. However, like Group 3, the correlation between sensitivity and discriminability was not significant ( $r = .30$ ,  $n = 36$ ,  $p = .08$ , two-tailed).

The above analyses calculated sensitivity by examining the change in response bias as a function of the relative frequency of punishment ( $\log P_{21}/P_{12}$ ). However, changes in response bias can also be assessed as a function of the combined effects of the reinforcers and

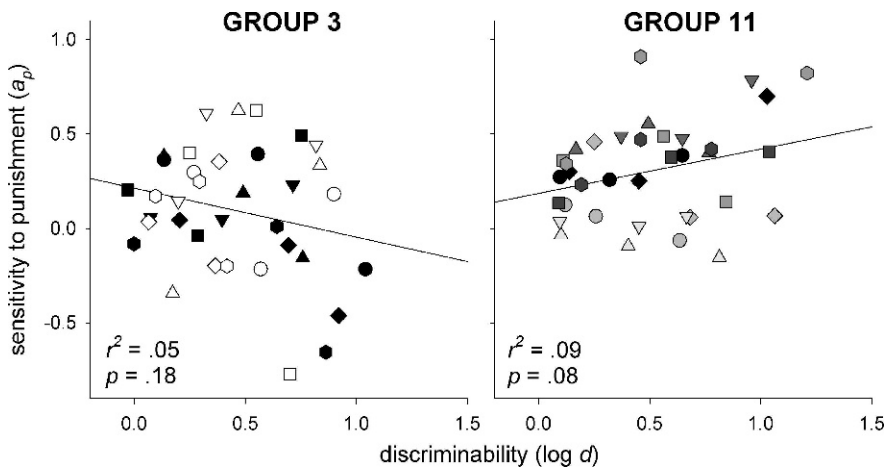


Fig. 5. Individual estimates of sensitivity to changes in punisher ratio ( $a_p$ ) are plotted as a function of discriminability ( $\log d$ ) for each participant in Group 3 (left) and Group 11 (right).

punishers. Researchers using standard concurrent-schedule choice procedures have proposed two competing models that attempt to combine reinforcer and punisher effects. An additive model of punishment (e.g., Deluty, 1976) predicts that the effects of punishers obtained for responding on one alternative add to the effects of reinforcers obtained for responding on the other alternative. On the other hand, a subtractive model (e.g., de Villiers, 1980; Farley & Fantino, 1978) predicts that the effects of punishers directly subtract from the effects of reinforcers for the same alternative. Lie and Alsop (2007, 2009) integrated each of the two competing models into Davison and Tustin’s (1978) model of signal detection. For the additive model, response bias can be calculated by

$$\begin{aligned} \log b &= 0.5 \log \left( \frac{B_{11} B_{21}}{B_{12} B_{22}} \right) \\ &= a_{r+p} \log \left( \frac{R_{11} + qP_{12}}{R_{22} + qP_{21}} \right) + \log c \end{aligned} \tag{7}$$

with all notation as above, and  $a_{r+p}$  is the sensitivity of the subject’s behavior to the combined additive effects of reinforcers and punishers. The scaling parameter  $q$  is used to equate the value of a punisher to the value of a reinforcer (e.g., if  $q = .5$ , then one punisher has half the subjective weighting of one reinforcer). For the subtractive model, response bias can be calculated by

$$\begin{aligned} \log b &= 0.5 \log \left( \frac{B_{11} B_{21}}{B_{12} B_{22}} \right) \\ &= a_{r-p} \log \left( \frac{R_{11} - qP_{21}}{R_{22} - qP_{12}} \right) + \log c \end{aligned} \tag{8}$$

with all notation as above, and  $a_{r-p}$  is the sensitivity of the subject’s behavior to the combined subtractive effects of reinforcers and punishers. Research on the combined effects of reinforcers and punishers using concurrent-schedule and signal-detection procedures has found stronger support for a subtractive model over an additive model (Critchfield, Paletz, MacAleese, & Newland, 2003; de Villiers, 1980; Farley, 1980; Farley & Fantino, 1978; Lie & Alsop, 2009).

Table 2 displays individuals’ estimates of sensitivity calculated by the additive model ( $a_{r+p}$ ) and the subtractive model ( $a_{r-p}$ ) for the

Table 2

Sensitivity to the combined effects of reinforcers and punishers, calculated using the additive model ( $a_{r+p}$  – Equation 7) and the subtractive model ( $a_{r-p}$  – Equation 8) for each participant in Group 3 and Group 11 at each stimulus disparity level.

Part.	Additive ( $a_{r+p}$ )			Subtractive ( $a_{r-p}$ )		
	Low	Med	High	Low	Med	High
GROUP 3						
1	1.31	1.85	−1.47	0.51	0.94	−0.93
2	1.21	−0.66	2.07	0.54	−0.21	0.97
3	0.25	0.16	1.00	0.11	0.05	0.51
4	−1.54	2.71	2.03	−0.69	1.18	1.15
5	0.68	−0.13	2.18	0.25	−0.04	0.99
6	1.80	3.06	−3.44	0.89	1.25	−1.59
7	0.19	−0.60	−3.13	0.10	−0.42	−1.95
8	0.13	−0.61	1.18	0.05	−0.15	0.45
9	1.25	0.71	−0.63	0.51	0.28	−0.35
10	0.60	2.25	2.24	0.26	0.90	1.09
11	−0.26	0.05	−2.86	−0.09	0.02	−1.45
12	0.63	0.76	−0.60	0.23	0.13	— <sup>a</sup>
GROUP 11						
13	1.40	1.13	2.98	0.62	0.34	1.59
14	0.70	0.28	−0.26	0.28	0.08	−0.07
15	2.78	6.31	7.24	1.30	3.91	4.29
16	−0.18	−0.40	−0.79	−0.07	−0.11	−0.30
17	0.79	2.94	5.14	0.36	1.84	3.99
18	2.09	3.66	0.81	1.06	2.23	0.46
19	1.59	1.20	6.68	0.65	0.41	3.79
20	2.38	0.48	0.62	1.10	0.28	0.41
21	2.16	3.31	2.73	0.87	1.24	1.67
22	0.15	0.05	0.36	0.05	0.01	0.15
23	1.27	2.72	3.04	0.54	1.35	1.88
24	1.81	8.06	10.58	0.77	5.35	8.42

<sup>a</sup> Participant 12 received more  $P_{21}$  punishers than  $R_{11}$  reinforcers in this condition so  $a_{r-p}$  could not be calculated.

three different levels of disparity for Group 3 and Group 11. For all calculations, it was assumed that  $q = 1$  because reinforcers and punishers were equivalent in magnitude (i.e., one-point gains and losses respectively). In all cases, estimates of sensitivity calculated by the additive model (Group 3: Low:  $M = 0.52$ , Medium:  $M = 0.80$ , High:  $M = −0.12$ ; Group 11: Low:  $M = 1.41$ , Medium:  $M = 2.48$ , High:  $M = 3.26$ ) were more extreme than the corresponding estimates calculated by the subtractive model (Group 3: Low:  $M = 0.22$ , Medium:  $M = 0.33$ , High:  $M = −0.10$ ; Group 11: Low:  $M = 0.63$ , Medium:  $M = 1.41$ , High:  $M = 2.19$ ). For both Group 3 and Group 11, sensitivity estimates spanned a wide range of values (Additive:  $−3.44$  to  $10.58$ ; Subtractive:  $−1.95$  to  $8.42$ ).

As before, we also examined whether sensitivity estimates significantly differed from zero.

For Group 3, estimates of sensitivity did not significantly differ from zero for each disparity level for either sensitivity measure (Additive: Low:  $t(11) = 2.052, p = .07$ ; Medium:  $t(11) = 2.050, p = .07$ ; High:  $t(11) = .188, p = .85$ ; Subtractive: Low:  $t(11) = 1.960, p = .08$ ; Medium:  $t(11) = 1.955, p = .08$ , High:  $t(10) = .281, p = .79$ ). For Group 11 however, estimates of sensitivity were significantly greater than zero at all disparity levels for both sensitivity measures (Additive: Low:  $t(11) = 5.389, p < .001$ ; Medium:  $t(11) = 3.295, p < .01$ ; High:  $t(11) = 3.222, p < .01$ ; Subtractive: Low:  $t(11) = 5.147, p < .001$ ; Medium:  $t(11) = 2.860, p < .05$ , High:  $t(11) = 2.978, p < .05$ ).

Finally, we examined whether the additive and subtractive sensitivity estimates significantly differed across the three disparity levels. For the additive model, a repeated-measures ANOVA found no significant differences in sensitivity estimates ( $a_{r+p}$ ) across the three disparity levels for Group 3 participants,  $F(2,22) = .972, p = .39$ , or Group 11 participants,  $F(2,22) = 3.213, p = .06$ . For the subtractive model however, the data violated the assumption of sphericity (Mauchly's  $W = .489, p < .05$ ). Using a Greenhouse-Geisser correction, the difference in sensitivity estimates ( $a_{r-p}$ ) between disparity levels approached significance,  $F(1.324, 14.56) = 4.210, p = .050$ . A cursory examination of the data found that sensitivity estimates appeared to increase as the stimulus disparity level increased (i.e., from low to high). To investigate this further, a Page's trend test was conducted; however, this test found no significant increase in  $a_{r-p}$  estimates across increases in stimulus disparity ( $L = 151, n = 12, k = 3$ ).

## DISCUSSION

The present study examined the relation between stimulus control and punisher control using human participants in a signal-detection task. There were three main findings. First, the participants' behavior was found to be under stimulus control. This was demonstrated by the significant effect that varying stimulus disparity had on the participants' ability to discriminate between the stimuli ('more blue' versus 'more red'). Participants made more correct responses when stimuli were more disparate than when they were less disparate (Figure 3, top). This

was consistent with Johnstone and Alsop (2000—Figure 2, top) who also found an increase in discriminability with increases in stimulus disparity using human participants. While this finding was not surprising, it was important to confirm that the disparity levels were sufficient to affect participants' estimates of discriminability.

Second, behavior was also under punisher control. Overall, participants were biased away from the response alternative associated with the higher frequency of punishment for each punisher ratio. The size of this effect was also somewhat dependent on the punisher ratio in place—larger response biases were found for Group 11 (1:11 and 11:1 punisher ratios), compared to Group 3 (1:3 and 3:1 punisher ratios). Significant differences in response bias were found between the two punisher ratios for both Groups 3 and 11. Furthermore, when estimates of sensitivity to the punisher ratios ( $a_p$ ) were calculated for each participant, sensitivity values were significantly greater than zero across all three disparity levels for Group 11, and the lowest disparity level for Group 3. Mean sensitivity estimates ( $a_p$ ) for Group 11 (between .25 and .32) were similar to the mean sensitivity estimate of .20 found by Lie and Alsop (2009), who used a similar task and sample of participants. Although these sensitivity to punishment ( $a_p$ ) values were slightly lower than sensitivity to reinforcement ( $a$ ) values found with human participants in previous studies using reinforcer manipulations (e.g., Alsop, Rowley, & Fon, 1995; Johnstone & Alsop, 2000; Lie & Alsop, 2009—mean sensitivity between .30 and .40), the effects of the punishers in the present experiment were probably attenuated by the effects of the constant and equal (1:1) distribution of reinforcers. However, this background rate of reinforcement was necessary to maintain participant responding.

Finally, the present experiment found no evidence of a relation between stimulus control and punisher control, consistent with the predictions of the Davison and Tustin (1978) model of signal detection. First, estimates of discriminability did not differ across the two punisher ratios for Group 3 or Group 11 and the mean estimates of discriminability were similar across groups at each level of disparity (Figure 3, top). Second, estimates of response bias (away from the more frequently

punished alternative) did not significantly differ across the three stimulus disparity levels for either group (Figure 3, bottom). This parallels the results of Johnstone and Alsop's (2000) study where they found no changes in response bias (towards the more frequently reinforced alternative) across their four disparity levels (Figure 2). Johnstone and Alsop also used a similar sample of participants (university students), similar stimuli (visual arrays), and the same reinforcer type (points) as the present study. Last, estimates of sensitivity to the punisher ratios ( $a_p$ ) were not significantly correlated with discriminability for either group in the present study (Figure 5). While this was consistent with nonhuman research looking at sensitivity to reinforcer ratios ( $a$ ) conducted by McCarthy and Davison (1984—Figure 1, top), it was inconsistent with the research conducted by Alsop and Davison (1991—Figure 1, middle) and Alsop and Porritt (2006—Figure 1, bottom).

Although the present experiment found no relation between stimulus control and punisher control, this finding should perhaps be interpreted with some caution because of the limited number of conditions that participants received; this was unfortunately due to time constraints on subject participation. It is possible that the addition of more punisher ratios may have revealed some interaction between the parameters of interest. For example, Johnstone and Alsop (1999) found lower estimates of discriminability ( $\log d$ ) for equal reinforcer ratios compared to unequal reinforcer ratios with nonhuman subjects (pigeons). It is unclear whether this would also have occurred had an equal punisher ratio (1:1) been included in the present experiment. However, Lie and Alsop (2009) arranged four punisher ratios (5:1, 2:1, 1:2, 1:5) in a similar procedure from the same participant pool as the present study and found no significant differences in discriminability across the four ratios. This perhaps suggests that the inclusion of additional punisher ratios in the present study would have yielded similar results to Lie and Alsop (2009).

Another drawback from the limited condition numbers in the present study was that estimates of sensitivity to the punisher ratios ( $a_p$ ) could only be calculated across two punisher ratios for each group. The present study was designed to be similar to Johnstone

and Alsop's (2000) study, where only between-group comparisons could be made between their two reinforcer ratios across different stimulus disparity levels. However, participants within each group in the present study experienced both punisher ratios, thus making it possible to evaluate within-subject changes in response bias between the two punisher ratios. By doing so in a standard way (i.e., calculation of sensitivity to punishment estimates), this allowed us to make tentative comparisons between the results of the present study to those of previous studies. Because of the limited condition numbers however, sensitivity estimates obtained from the present study should be interpreted with caution.

More generally, it is unclear why some studies and not others have found a relation between sensitivity (or response bias) and discriminability. The present study, and those mentioned above (Alsop & Davison, 1991; Alsop & Porritt, 2006; Johnstone & Alsop, 2000; McCarthy & Davison, 1984) all employed controlled reinforcer procedures (i.e., dependent or interdependent scheduling), and have obtained mixed results. Other detection studies using controlled reinforcer procedures have found decreases in sensitivity with increases in discriminability (Davison & McCarthy, 1987; Godfrey & Davison, 1998; Nevin, Cate, & Alsop, 1993), but also no relation between sensitivity and discriminability (Godfrey & Davison, 1998). Studies which have arranged reinforcers using uncontrolled procedures (where the obtained reinforcer ratios can co-vary with the subject's behavior, i.e., independent scheduling) have also found conflicting findings. Again, some of these studies have found decreases in sensitivity with increases in discriminability (Johnstone & Alsop, 2000; McCarthy & Davison, 1984) or no such relation between the two parameters (e.g., McCarthy & Davison, 1980). Thus, although the present experiment found no relation between sensitivity and discriminability using a controlled punisher ratio procedure, it seems that procedural variations can influence the findings obtained. In fact, the task of designing an experiment to adequately test the relation between discriminability and response bias appears to be quite a challenge. As described by Nevin (1984), any such experiment needs to arrange suitable conditions which demonstrate that: (1)  $\log d$  does

not vary systematically across changes in the reinforcer (or punisher) ratio; (2) increases in stimulus disparity lead to consistent increases in  $\log d$ ; and (3) that the effects of varied reinforcer (or punisher) ratios (i.e.,  $\log b$ ,  $a$ , or  $a_p$ ) are unchanged by variations in  $\log d$ . Furthermore, these effects need to be demonstrated across a number of stimulus modalities as well as species. Given the mixed findings described above, it appears that this overall approach has not yet been achieved.

The present experiment also calculated sensitivity by using the modified Davison and Tustin (1978) equations proposed by Lie and Alsup (2007; 2009). Sensitivity to the combined additive ( $a_{r+p}$ ) and subtractive ( $a_{r-p}$ ) effects of reinforcement and punishment were calculated using Equations 7 and 8; however, these were not particularly sensible. In fact, the ranges of sensitivity estimates (Additive:  $-3.44$  to  $10.58$ ; Subtractive:  $-1.95$  to  $8.42$ ) were well outside the usual range of sensitivity to reinforcement estimates ( $a$ ) obtained from human detection experiments which studied the effects of reinforcers alone (e.g.,  $-0.01$  to  $0.58$ ; Lie & Alsup, 2009). These additive and subtractive model fits were conducted under the assumption that  $q$  (the scaling parameter) was equal to 1 for all participants. While this assumption appeared reasonable because the reinforcers and punishers were equal in physical value (i.e., one point gain vs. one point loss), it is possible that their subjective values differed across participants. Had we allowed  $q$  to vary as well as  $a_{r+p}$  for the additive model (or  $a_{r-p}$  for the subtractive model), more sensible sensitivity estimates might have been obtained. However, with only two conditions in the present experiment (and two free parameters), this was not possible. Nevertheless, this was the first attempt to fit the additive and subtractive models to empirical data. Future experiments arranging more conditions could explore the quantitative predictions of the additive and subtractive punishment models further.

Although the integration of punisher effects into Davison and Tustin's (1978) well known behavioral model appears to be the most logical first step, Alsup and Davison (1991; see also Davison & Nevin, 1999) have also proposed a competing detection model based on the discriminability of stimulus-response ( $d_s$ ) and response-reinforcer ( $d_r$ ) relations.

The Alsup and Davison model has been somewhat successful in capturing the effects of reinforcers independent of stimulus discriminability (Davison & Nevin, 1999; Godfrey & Davison, 1998). However, as noted by Lie and Alsup (2009), it is unclear how the Alsup and Davison model can be extended to include punisher as well as reinforcer effects. For example, would separate stimulus-response discriminability and response-reinforcer discriminability parameters be required for reinforcers and punishers? Because the present experiment found no relation between punisher effects ( $\log b$  and  $a$ ) and stimulus effects ( $\log d$ ) using the Davison-Tustin model, and punisher effects are more easily integrated into the Davison-Tustin model than the Alsup-Davison model, it seems more parsimonious to use the former instead of the latter at this stage.

Finally, the present experiment was not without limitations. Although an effort was made to equate the overall reinforcer and punisher rates across stimulus disparity levels, this proved to be quite difficult, especially in the high disparity conditions. In fact, significant differences in overall reinforcer rates were found across the three stimulus disparity levels for Group 3,  $F(2,46) = 29.12$ ,  $p < .001$ , and Group 11,  $F(2,26) = 19.98$ ,  $p < .001$ . Similarly, significant differences in overall punisher rates were also found, Group 3:  $F(2,46) = 10.96$ ,  $p < .001$ ; Group 11:  $F(2,46) = 9.399$ ,  $p < .001$ . While previous research has found that changes in overall reinforcer rates can affect sensitivity to reinforcement in concurrent-schedule procedures (Alsup & Elliffe, 1988), a wide range of reinforcer rates was used to demonstrate the effect (i.e., from 0.22 to 10 reinforcers per min). In the present experiment, mean reinforcer rates for each disparity level ranged from 2.3 to 2.7 reinforcers per min and mean punisher rates ranged from 0.60 to 1.08 punishers per min. Thus, it seems quite unlikely that such small changes in overall reinforcer and punisher rates would affect sensitivity estimates.

Because of the relatively low rates of punishment, there were 8 sessions (out of a total of 144) where a few participants received no punishers for responding to one response alternative (see Appendix). In all eight cases, participants showed a bias towards responding on the alternative where they received no



punishers; this may demonstrate that very few punishers are necessary to influence choice behavior in this task. Upon closer inspection, seven of the eight sessions were high disparity conditions where participants were very accurate (discriminability  $M = 1.01$ ), and thus made few errors. This is a difficulty inherent in arranging high disparity conditions in detection procedures that punish incorrect responses. Future studies on the effects of punishment in detection procedures should take this into consideration and either arrange very many trials per condition, or make stimuli less disparate in their highest disparity condition(s) so that participants come into adequate contact with the punisher contingencies.

The present experiment is the first to examine the relation between punisher control and stimulus control in signal-detection procedures. While studies of signal detection have largely focused on the effects of positive outcomes for correct responses, it is also important to study the effects of negative outcomes for errors because organisms encounter both types of outcomes in many everyday situations. The present study thus provides a possible direction for future research on the effects of negative outcomes on human (as well as nonhuman) behavior in situations of uncertainty.

## REFERENCES

- Alsop, B., & Davison, M. (1991). Effects of varying stimulus disparity and the reinforcer ratio in concurrent-schedule and signal-detection procedures. *Journal of the Experimental Analysis of Behavior*, 56, 67–80.
- Alsop, B., & Elliffe, D. (1988). Concurrent-schedule performance: Effects of relative and overall reinforcer rate. *Journal of the Experimental Analysis of Behavior*, 49, 21–36.
- Alsop, B., & Porritt, M. (2006). Discriminability and sensitivity to reinforcer magnitude in a detection task. *Journal of the Experimental Analysis of Behavior*, 85, 41–56.
- Alsop, B., Rowley, R., & Fon, C. (1995). Human symbolic matching-to-sample performance: Effects of reinforcer and sample-stimulus probabilities. *Journal of the Experimental Analysis of Behavior*, 63, 53–70.
- Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22, 231–242.
- Brown, G. S., & White, K. G. (2005). The optimal correction for estimating extreme discriminability. *Behavior Research Methods*, 37, 436–449.
- Critchfield, T. S., Paletz, E. M., MacAleese, K. R., & Newland, M. (2003). Punishment in human choice: Direct or competitive suppression? *Journal of the Experimental Analysis of Behavior*, 80, 1–27.
- Davison, M., & McCarthy, D. (1987). The interaction of stimulus and reinforcer control in complex temporal discrimination. *Journal of the Experimental Analysis of Behavior*, 48, 97–116.
- Davison, M., & Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: An integration. *Journal of the Experimental Analysis of Behavior*, 71, 439–482.
- Davison, M., & Tustin, R. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of the Experimental Analysis of Behavior*, 29, 331–336.
- Deluty, M. Z. (1976). Choice and the rate of punishment in concurrent schedules. *Journal of the Experimental Analysis of Behavior*, 25, 75–80.
- de Villiers, P. (1980). Toward a quantitative theory of punishment. *Journal of the Experimental Analysis of Behavior*, 33, 15–25.
- Farley, J. (1980). Reinforcement and punishment effects in concurrent schedules: A test of two models. *Journal of the Experimental Analysis of Behavior*, 33, 311–326.
- Farley, J., & Fantino, E. (1978). The symmetrical law of effect and the matching relation in choice behavior. *Journal of the Experimental Analysis of Behavior*, 29, 37–60.
- Godfrey, R., & Davison, M. (1998). Effects of varying sample- and choice-stimulus disparity on symbolic matching-to-sample performance. *Journal of the Experimental Analysis of Behavior*, 69, 311–326.
- Hume, A. L., & Irwin, R. J. (1974). Bias functions and operating characteristics of rats discriminating auditory stimuli. *Journal of the Experimental Analysis of Behavior*, 21, 285–295.
- Johnstone, V., & Alsop, B. (1999). Stimulus presentation ratios and the outcomes for correct responses in signal-detection procedures. *Journal of the Experimental Analysis of Behavior*, 72, 1–20.
- Johnstone, V., & Alsop, B. (2000). Reinforcer control and human signal-detection performance. *Journal of the Experimental Analysis of Behavior*, 73, 275–290.
- Lie, C., & Alsop, B. (2007). Reinforcement and punishment in behavioral models of signal detection. *Revista Mexicana de Analisis de la Conducta*, 33, 45–55.
- Lie, C., & Alsop, B. (2009). Effects of point-loss punishers on human signal-detection performance. *Journal of the Experimental Analysis of Behavior*, 92, 17–39.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- McCarthy, D., & Davison, M. (1979). Signal probability, reinforcement and signal detection. *Journal of the Experimental Analysis of Behavior*, 32, 373–386.
- McCarthy, D., & Davison, M. (1980). Independence of sensitivity to relative reinforcement rate and discriminability in signal detection. *Journal of the Experimental Analysis of Behavior*, 34, 273–284.
- McCarthy, D., & Davison, M. (1984). Isobias and alloibias functions in animal psychophysics. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 390–409.
- Nevin, J. A. (1984). Quantitative analysis. *Journal of the Experimental Analysis of Behavior*, 42, 421–434.

- Nevin, J. A., Cate, H., & Alsop, B. (1993). Effects of differences between stimuli, responses, and reinforcer rates on conditional discrimination performance. *Journal of the Experimental Analysis of Behavior*, 59, 147–161.
- Stubbs, D. A., & Pliskoff, S. S. (1969). Concurrent responding with fixed relative rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, 12, 887–895.
- Terman, M. (1970). Discrimination of auditory intensities by rats. *Journal of the Experimental Analysis of Behavior*, 13, 145–160.

Received: July 1, 2009

Final Acceptance: December 30, 2009

## APPENDIX

The numbers of reinforcers ( $R_{11}$  and  $R_{22}$ ) and punishers ( $P_{21}$  and  $P_{12}$ ) calculated across all trials, and the numbers of  $B_{11}$ ,  $B_{12}$ ,  $B_{21}$ , and  $B_{22}$  responses, estimates of discriminability ( $\log d$ ) and response bias ( $\log b$ ) calculated across the last 120 trials, for each participant (Ppt.) in each condition (Con.) in Group 3 and Group 11. Condition order and the total numbers of trials per condition are also presented.

Ppt.	Con.	Order	Trials	$R_{11}$	$R_{22}$	$P_{21}$	$P_{12}$	$B_{11}$	$B_{12}$	$B_{21}$	$B_{22}$	$\log d$	$\log b$
<b>GROUP 3</b>													
1	3:1H	5	266	30	28	6	2	52	7	5	56	0.94	−0.09
	1:3H	2	283	34	32	4	12	52	7	2	59	1.14	−0.28
	3:1M	6	303	39	37	19	7	43	17	12	48	0.50	−0.10
	1:3M	3	261	36	36	6	16	51	7	18	44	0.62	0.23
	3:1L	4	395	42	43	26	9	39	22	21	38	0.25	0.00
2	1:3L	1	431	44	45	9	30	42	18	41	19	0.02	0.35
	3:1H	2	365	42	41	25	8	56	4	16	44	0.78	0.34
	1:3H	5	244	26	28	0	4	59	1	18	42	1.02	0.66
	3:1M	1	391	46	48	35	11	48	12	20	40	0.45	0.15
	1:3M	4	412	48	46	11	33	48	11	9	52	0.69	−0.06
3	3:1L	3	421	42	42	26	8	31	29	20	40	0.16	−0.13
	1:3L	6	323	36	36	6	16	46	14	22	38	0.37	0.14
	3:1H	1	376	42	40	24	9	49	12	8	51	0.70	−0.09
	1:3H	4	296	34	31	3	12	53	7	12	48	0.73	0.14
	3:1M	3	525	57	58	48	17	46	15	23	36	0.34	0.14
4	1:3M	6	339	40	40	8	23	49	11	21	39	0.45	0.19
	3:1L	2	382	42	39	24	7	35	24	29	32	0.10	0.06
	1:3L	5	430	42	41	9	24	35	24	33	28	0.05	0.12
	3:1H	2	273	29	28	5	2	41	19	3	57	0.79	−0.46
	1:3H	5	368	36	35	4	17	52	9	5	54	0.88	−0.13
5	3:1M	3	334	30	30	8	3	20	40	4	56	0.41	−0.71
	1:3M	6	405	45	44	10	29	41	18	10	51	0.53	−0.17
	3:1L	1	414	42	41	24	9	27	33	17	43	0.16	−0.24
	1:3L	4	366	36	38	6	19	18	42	9	51	0.19	−0.55
	3:1H	4	319	33	32	11	4	35	25	2	58	0.78	−0.63
6	1:3H	1	410	43	42	9	26	47	14	6	53	0.73	−0.20
	3:1M	6	570	54	54	45	14	34	27	17	42	0.24	−0.14
	1:3M	3	352	44	44	10	29	35	25	14	46	0.33	−0.18
	3:1L	5	520	44	46	31	9	16	43	20	41	−0.06	−0.37
	1:3L	2	466	45	45	11	29	24	36	24	36	0.00	−0.17
7	3:1H	6	373	43	44	27	10	50	9	24	37	0.46	0.27
	1:3H	3	276	32	32	3	11	45	15	2	58	0.94	−0.47
	3:1M	4	269	32	30	9	3	35	26	4	55	0.62	−0.49
	1:3M	1	353	44	43	10	29	47	13	17	43	0.48	0.08
	3:1L	5	390	37	36	17	6	27	33	11	49	0.28	−0.36
8	1:3L	2	327	38	38	6	20	38	22	23	37	0.22	0.02
	3:1H	1	301	33	33	12	4	58	1	26	35	0.90	0.77
	1:3H	4	260	28	29	3	5	57	2	15	46	0.94	0.46
	3:1M	2	240	31	30	9	3	56	5	23	36	0.61	0.42
	1:3M	5	212	30	29	3	6	57	4	16	43	0.78	0.35
9	3:1L	3	369	36	37	18	6	49	12	33	26	0.25	0.35
	1:3L	6	319	36	37	6	17	47	13	38	22	0.16	0.39

APPENDIX

(Continued)

Ppt.	Con.	Order	Trials	R <sub>11</sub>	R <sub>22</sub>	P <sub>21</sub>	P <sub>12</sub>	B <sub>11</sub>	B <sub>12</sub>	B <sub>21</sub>	B <sub>22</sub>	log <i>d</i>	log <i>b</i>
8	3:1H	1	423	41	42	25	9	42	18	11	49	0.50	−0.14
	1:3H	4	464	45	44	11	29	43	16	27	34	0.26	0.16
	3:1M	2	560	58	58	51	16	50	10	19	41	0.51	0.18
	1:3M	5	441	49	49	12	37	37	23	22	38	0.22	−0.02
	3:1L	3	514	46	45	30	11	32	29	27	32	0.06	−0.02
9	1:3L	6	349	39	38	7	20	33	27	28	32	0.07	0.01
	3:1H	5	363	36	37	17	6	50	9	13	48	0.65	0.09
	1:3H	2	282	35	32	4	13	53	8	6	53	0.87	−0.06
	3:1M	4	279	36	37	17	6	46	14	11	49	0.58	−0.06
	1:3M	1	396	48	50	12	36	46	14	20	40	0.40	0.11
10	3:1L	6	366	36	40	20	6	28	31	14	47	0.24	−0.28
	1:3L	3	385	46	44	9	31	35	25	33	27	0.03	0.12
	3:1H	6	241	28	28	4	2	51	11	3	55	0.94	−0.29
	1:3H	3	432	39	39	7	22	51	9	11	49	0.69	0.05
	3:1M	5	356	41	42	25	8	28	32	10	50	0.32	−0.37
11	1:3M	2	353	43	44	9	28	47	13	26	34	0.33	0.22
	3:1L	4	432	46	42	29	9	35	25	21	39	0.21	−0.06
	1:3L	1	347	39	38	6	21	40	21	26	33	0.19	0.09
	3:1H	4	332	34	35	15	5	61	1	17	41	1.03	0.66
	1:3H	1	398	38	39	7	20	52	9	11	48	0.69	0.06
12	3:1M	5	356	47	43	31	9	52	8	24	36	0.49	0.31
	1:3M	2	324	36	38	6	18	56	4	15	45	0.80	0.32
	3:1L	6	484	48	48	35	11	38	22	38	22	0.00	0.24
	1:3L	3	483	45	42	10	27	36	25	35	24	0.00	0.16
	3:1H	3	594	56	56	57	19	42	17	21	40	0.33	0.06
13	1:3H	6	480	48	48	11	35	42	18	11	49	0.50	−0.14
	3:1M	2	652	55	55	54	17	35	23	23	39	0.20	−0.02
	1:3M	5	392	44	43	9	28	49	12	24	35	0.38	0.22
	3:1L	1	486	45	45	31	9	35	26	26	33	0.12	0.01
	1:3L	4	652	44	45	9	30	39	21	34	26	0.08	0.19
GROUP 11													
14	11:1H	2	291	26	29	5	0	31	29	1	59	0.85	−0.82
	1:11H	5	392	38	39	2	26	46	13	19	42	0.44	0.10
	11:1M	1	546	48	48	43	4	23	37	11	49	0.22	−0.42
	1:11M	4	354	42	42	3	31	46	14	19	41	0.42	0.09
	11:1L	3	440	35	36	20	2	17	43	13	47	0.08	−0.48
15	1:11L	6	447	41	40	2	29	37	23	29	31	0.12	0.09
	11:1H	6	461	45	47	40	4	51	10	10	49	0.69	0.01
	1:11H	3	474	49	50	4	45	44	15	10	51	0.58	−0.12
	11:1M	5	498	53	51	49	5	36	26	15	43	0.30	−0.16
	1:11M	2	381	44	44	3	35	36	23	22	39	0.22	−0.03
16	11:1L	4	416	42	39	29	2	31	30	21	38	0.13	−0.12
	1:11L	1	381	41	42	2	31	39	21	32	28	0.10	0.16
	11:1H	4	277	31	35	15	1	47	12	6	55	0.77	−0.18
	1:11H	1	272	25	26	0	1	59	0	33	28	1.15	1.22
	11:1M	6	308	27	25	4	0	18	43	1	58	0.65	−1.02
17	1:11M	3	253	33	34	1	16	51	9	13	47	0.65	0.10
	11:1L	5	393	38	37	23	2	14	45	2	59	0.46	−0.96
	1:11L	2	359	38	39	2	25	41	19	22	38	0.28	0.05
	11:1H	6	313	34	34	16	2	52	8	10	50	0.75	0.06
	1:11H	3	406	45	45	3	37	49	11	4	56	0.88	−0.24
18	11:1M	4	320	40	41	29	3	36	23	13	48	0.38	−0.18
	1:11M	1	434	51	51	4	48	32	28	8	52	0.43	−0.37
	11:1L	5	383	40	39	27	2	22	38	19	41	0.05	−0.28
	1:11L	2	423	44	45	3	36	23	37	14	46	0.15	−0.36

APPENDIX

(Continued)

Ppt.	Con.	Order	Trials	R <sub>11</sub>	R <sub>22</sub>	P <sub>21</sub>	P <sub>12</sub>	B <sub>11</sub>	B <sub>12</sub>	B <sub>21</sub>	B <sub>22</sub>	log <i>d</i>	log <i>b</i>
17	11:1H	5	268	27	27	4	0	43	18	1	58	1.02	-0.65
	1:11H	2	279	30	29	1	8	57	3	7	53	1.06	0.19
	11:1M	6	426	34	31	14	1	24	37	4	55	0.46	-0.65
	1:11M	3	254	33	31	1	13	53	7	12	48	0.73	0.14
	11:1L	4	332	37	37	22	2	28	33	23	36	0.06	-0.13
	1:11L	1	362	38	42	2	28	39	21	31	29	0.12	0.15
18	11:1H	1	302	33	35	17	2	47	13	5	55	0.79	-0.23
	1:11H	4	321	34	32	1	15	55	6	7	52	0.90	0.04
	11:1M	2	372	30	30	10	1	26	33	3	58	0.58	-0.68
	1:11M	5	293	34	34	1	17	54	7	22	37	0.55	0.33
	11:1L	3	432	33	35	17	1	14	47	10	49	0.08	-0.60
	1:11L	6	382	38	38	2	24	40	19	32	29	0.14	0.18
19	11:1H	3	252	26	25	1	0	31	28	0	61	1.22	-1.17
	1:11H	6	324	35	33	2	16	53	9	6	52	0.84	-0.08
	11:1M	1	347	46	45	39	3	41	20	17	42	0.35	-0.04
	1:11M	4	351	40	41	3	28	55	5	27	33	0.55	0.47
	11:1L	2	348	33	34	16	1	18	42	8	52	0.22	-0.58
	1:11L	5	487	45	45	4	36	33	27	29	31	0.06	0.03
20	11:1H	5	308	33	34	15	2	60	1	7	52	1.27	0.41
	1:11H	2	221	28	27	0	5	58	2	20	40	0.86	0.56
	11:1M	4	257	34	33	16	1	48	12	8	52	0.70	-0.10
	1:11M	1	241	31	33	1	14	50	10	11	49	0.67	0.02
	11:1L	6	394	37	37	23	2	19	43	5	53	0.33	-0.68
	1:11L	3	325	39	37	2	24	44	16	33	27	0.17	0.26
21	11:1H	1	238	30	31	11	1	43	17	5	55	0.71	-0.31
	1:11H	4	330	33	34	2	16	56	3	17	44	0.82	0.41
	11:1M	3	304	32	31	12	1	27	32	3	58	0.59	-0.66
	1:11M	6	494	43	46	4	38	53	8	30	29	0.40	0.41
	11:1L	2	354	43	42	32	3	24	36	12	48	0.21	-0.38
	1:11L	5	425	38	39	2	26	49	12	41	18	0.13	0.48
22	11:1H	3	299	30	31	11	1	44	17	7	52	0.63	-0.22
	1:11H	6	416	43	44	4	34	48	12	8	52	0.70	-0.10
	11:1M	2	381	42	42	31	3	40	19	12	49	0.46	-0.14
	1:11M	5	502	53	52	5	50	40	19	13	48	0.44	-0.12
	11:1L	1	479	46	45	37	4	28	32	22	38	0.09	-0.15
	1:11L	4	487	47	44	3	39	31	29	24	36	0.10	-0.07
23	11:1H	2	289	32	34	15	1	47	15	2	56	0.94	-0.45
	1:11H	5	310	31	32	1	12	56	5	23	36	0.61	0.42
	11:1M	3	313	32	32	12	2	32	29	3	56	0.64	-0.60
	1:11M	6	346	39	39	2	26	48	14	28	30	0.28	0.25
	11:1L	1	364	41	38	26	3	33	29	20	38	0.17	-0.11
	1:11L	4	435	40	41	2	29	46	12	36	26	0.22	0.36
24	11:1H	4	253	28	28	7	0	40	21	1	58	0.97	-0.70
	1:11H	1	243	28	27	1	4	61	0	14	45	1.45	0.94
	11:1M	5	448	31	30	12	1	15	46	2	57	0.46	-0.94
	1:11M	2	280	30	31	1	10	57	3	41	19	0.46	0.79
	11:1L	6	547	37	38	26	2	17	42	11	50	0.13	-0.52
	1:11L	3	460	38	42	3	27	40	21	31	28	0.12	0.16